Albert Gatt

Dep. Information and Computing Sciences, Utrecht University

Full Professor   May 2023 to present

Associate Professor (UHD) in NLP   August 2021 to April 2023

- I hold the Chair in Natural Language Generation (NLG) and co-lead the Natural Language Processing Group.
- Co-coordinator of the MSc. in Applied Data Science (ADS).
- Teaching and supervision at undergraduate and postgraduate (MSc) levels. Topics that I have taught include Research Methods for Computing Science and Pattern Recognition.

University of Malta

Associate Professor in Human Language Technology   2018 to 2021

Senior Lecturer in Human Language Technology   2012 to 2018

Lecturer in Human Language Technology   2007 to 2012

- Director of the Institute of Linguistics and Language Technology (LLT; 2012-2020), responsible for academic programme development and research development.
- Founder and leader of the cross-faculty research group in NLP.
- Researcher in Natural Language Processing, focusing on Natural Language Generation, the Vision-Language interface, and NLP for under-resourced languages.
- Teaching at undergraduate (BA/BSc) and postgraduate (MSc/Doctoral School) levels.
- Supervision at postgraduate (MSc/PhD) and undergraduate level.

Tilburg Center for Communication and Cognition (TiCC), Tilburg University, The Netherlands

Honorary Research Associate   2015 to 2018

Postdoctoral Research Fellow (0.2 fte)   2010 to 2013

- Visiting Research Fellow (2010-13), on the NWO VICI project *Bridging the gap between psycholinguistics and computational linguistics: The case of referring expressions*, led by prof. dr. Emiel Krahmer.
- Honorary Research Associate (2015-17), an honorary, unpaid, non-resident position in recognition of previous and ongoing research collaborations.

Department of Computing Science, University of Aberdeen, UK

Honorary Research Fellow   2009 to 2014
Postdoctoral Research Fellow   2007 to 2009

- As a post-doc, and subsequently as honorary RF, I was engaged on the BabyTalk Project, for which I developed core NLG technologies for automatically summarising multi-channel, heterogeneous, high-density data in Natural Language.
- I also collaborated on the EPSRC Platform Grant *Affecting People with Natural Language* with the other members of the Aberdeen NLG Group.
- My work was mainly under the supervision of prof. dr. Ehud Reiter, prof. dr. Jim Hunter and prof. dr. Kees van Deemter.

Faculty of Engineering, University of Malta

Research Assistant (0.5 fte)   2002

- I worked on a project funded by the Foundation for Information Technology Accessibility (Malta), aiming to develop Text to Speech technology for Maltese.
- I worked under the supervision of prof. dr. Paul Micallef.

Institute of Linguistics and Dept. Computer Science and AI, University of Malta

Research Assistant   2001 to 2002

- I was engaged as a full-time RA on the MaltiLex Project, which developed corpora and electronic lexical resources for Maltese.
- I worked under the supervision of prof. dr. Ray Fabri and mr. Michael Rosner.

———————————————

EDUCATION   **PhD Computing Science**   2004 to 2007
Information Technology Research Institute, University of Brighton (2004-2005)
Department of Computing Science, University of Aberdeen (2005-2007)
Thesis: *Generating Coherent References to Multiple Entities*.
Supervisors: prof. dr. Kees van Deemter; prof. dr. Ehud Reiter
Committee: prof. dr. Chris Mellish; prof. dr. Robert Dale
This PhD formed part of the EPSRC-funded TUNA Project.

M.A. in Linguistics (Formal Semantics)   2000 to 2002
University of Malta. Awarded *cum laude*.

B. Psychology (Hons.) with Linguistics   1996 to 2000
University of Malta. Awarded *cum laude*.

ESSLLI: European Summer School in Logic, Lanuage and Information 2001
University of Trento, Italy.

Exchange student, Pittsburgh University, PA, USA   1998
Courses in Psychology, Linguistics and Neuroscience. Inducted into *Phi Beta Kappa* in recognition of high academic achievement.

———————————————

| PROFESSIONAL DEVELOPMENT | **Senior Research Qualification (SKOz)** 2021 |
| | |

Senior Research Qualification (SKOz) 2021
Higher research qualification, awarded by Utrecht University in November, 2021.

Supervising PhD Research 2022
Course offered by the Continuing Education division at Utrecht University, aiming to expand the skills of already-experienced PhD supervisors to offer more efficient and effective supervision.

Academic Leadership for Associate Professors PhD Research 2022 to 2023
Course for recently appointed UHD members of Utrecht University, focusing on leadership roles and skills in an academic context. I was nominated by the Dean of the Faculty of Science to follow this course.

Senior Teaching Qualification 2022 to 2023
At Utrecht University, I have initiated the process to acquire my Senior Teaching Qualification, which I plan to complete by Spring 2023.

TEACHING EXPERIENCE

Undergraduate courses taught (selection) The following is a selection of Bachelor-level courses for which I was (co-)responsible for curriculum development and delivery.

- INFOB3OMI *Onderzoeksmethoden voor Informatica*, Utrecht University. Course in statistical analysis and general research methods for computer science undergraduates. Role: tutor.
- LIN3011 *Data-Driven Natural Language Processing*, University of Malta. Advanced undergraduate course for students following a BSc in Human Language Technology, BSc in Artificial Intelligence and BSc in Computer Science. Role: coordinator, sole lecturer.
- LIN1302 *Quantitative Approaches to Natural Language Analysis*, University of Malta. Undergraduate course in statistics for BA Linguistics and BSc Human Language Technology students. Role: coordinator, sole lecturer.
- LIN2302 *Experimental Techniques and Paradigms*, University of Malta. Undergraduate course in experimental design and advanced statistics. Role: coordinator, sole lecturer.

Postgraduate courses (selection) The following is a selection of courses at Master and Doctoral level for which I was (co-)responsible for curriculum development and delivery.

- INFOMPR *Pattern Recognition*, Utrecht University. Offered to MSc Artificial Intelligence and other MSc programs. Role: co-lecturer, project supervisor.
- LIN5507 *Social Media as Multimodal text*, University of Malta. Offered to M.Sc. Language and Computation and M.Sc. Human Language Science and Technology students. Role: coordinator, sole lecturer.
- CGS5080 *Artificial Intelligence*, University of Malta. Offered to M.Sc. Cognitive Science students. Role: co-lecturer.

ACADEMIC LEADERSHIP

Coordinator, MSc in Applied Data Science 2022 to present
Shortly after joining Utrecht University, I was selected to become a member of the coordination team of the Applied Data Science Masters program, which is co-led by the ICS Department with contributions from 4 other entities across the University.

In this role, I bring to bear some of my experience in interdisciplinary education management. One of my primary roles is to liaise with external agencies and companies, as well as UU staff across multiple departments, to identify projects for our students and coordinate the supervision, research and grading process.

Chair, Training and Career Development Committee, NL4XAI   2020 to present

NL4XAI is a Marie Skłodowska-Curie Innovative Training Network with 10 beneficiary institutions, including Utrecht University. Since its inception, I have chaired the committee responsible for developing training programs for our Early-Stage Researchers and monitoring the progress of their PhD projects.

Leader, NLP Research Group, University of Malta   2019 to 2021

The NLP Research Group brings together staff and research students from different faculties and centres of the University of Malta (UM), who collaborate in various projects. This group developed out of an earlier group called RiVaL (Research in Vision and Language), to accommodate a broader range of interests in NLP and AI. I founded this group to bring together researchers with common interests working in different faculties, to develop the critical mass required to conduct effective NLP research at UM.

Academic Program Coordinator, Institute of Linguistics and Langauge Technology   2009-2012

Coordinator of the undergraduate programmes, BA Linguistics and BSc. Human Language Technology.

Director, Institute of Linguistics and Language Technology (LLT)   2012 to 2020

- Primary responsibility for design and update of programmes of study at undergraduate and postgraduate level, and implementation of academic quality control measures. Responsible for articulating the academic vision of the LLT, diversifying its programmes in both computational and theoretical/applied areas of language.
- I also helped to forge research and teaching collaborations with other entities at the University of Malta, including the Faculty of ICT, the Faculty of Media and Knowledge Sciences and the Faculty of Arts.
- Additional roles:
    - Chair of the Doctoral Committee, Institute of Linguistics and Language Technology.
    - Member, Board of Studies, Erasmus Mundus Master of Science in Human Language Science and Technology.
    - Chair, Board of the Institute of Linguistics and Language Technology.
    - Member (by invitation), Faculty Board of the Faculty of Arts.

Chair, Creative Writing Programme, Faculty of Arts   2015 to 2019

- I took up the chairmanship of the CWP at the invitation of the Dean of the Faculty of Arts.
- I developed initiatives for interdisciplinary training for creative writers and organised seminars by leading international writers.

Appointments as visiting researcher
- June 2017: Visiting researcher, Centro de Investigación en Tecnologías de la Información y de las Comunicaciones de la Universidad de Granada (CITIC-UGR), Department of Computer Science and Artificial Intelligence, University of Granada, Spain.
- July - August 2015: Visiting researcher, Natural Language Generation Group, Department of Computing Science, University of Aberdeen, UK.

International teaching engagements (selection)
- May, 2023: *Multimodal deep models for NLP.* University of Malta. A 10-hour course for masters students of the European Master in Human Language Science and Technology.
- Octobter, 2022: *Topic Models and Embeddings for Information Retrieval.* Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence, University of Groningen, The Netherlands. Guest lecture in the course *Information Retrieval* (WBCS040-05, Bachelor level).
- September, 2022: *Four stories on visual grounding.* Lecture at the Multi3Generation PhD Summer School on *Representation-Mediated Multimodality.* Schloss Etelsen, Germany.
- January 2018: *Natural Language Generation.* Postgraduate-level course at the *LOT Winter School in Linguistics, Amsterdam, The Netherlands.* Co-taught with prof. dr. Emiel Krahmer (Tilburg University, The Netherlands).
- October 2015, May 2016, April 2018: *NLP Tools for Translation.* Invited seminars for professional translators. Directorate-General for Translation, the European Parliament, European Commission and European Council.
- July 2015: *Microplanning and realisation.* Lectures at the *PhD Summer School on Natural Language Generation, Summarisation and Dialogue Systems*, University of Aberdeen, UK.
- February 2014, March 2015, March 2016: *Statistics and Data Analysis for the Humanities and Social Sciences.* Intensive PhD course at the Doctoral School, University of Copenhagen, Denmark. Co-taught with prof. dr. Patrizia Paggio (CST, University of Copenhagen and University of Malta).
- August 2014: *Referring expression generation and multimodal reference.* Advanced PhD courses at the RefNet Summer School on Psychological and Computational Models of Language Production. School of Informatics, University of Edinburgh. Co-taught with dr. Ielka van der Sluis (University of Groningen, The Netherlands), dr. Paul Piwek (Open University, UK) and prof. dr. Kees van Deemter (formerly University of Aberdeen, now at the University of Utrecht).
- July 2012: *NLG and Data-to-Text.* PhD course at the HIT-MSRA Postgraduate Summer School on Human Language Technology. Harbin Institute of Technology, China.

Face:LIFT: Lifelike Images From Text 2021 to 2023

- Funder: Malta Council for Science and Technology, FUSION Programme
- Collaborators: Threls Ltd (industrial partner).
- Roles: PI (with Dr Reuben Farrugia, Department of Communications and Computer Engineering, University of Malta). Since August 2021, I am external collaborator on this project, following my move to Utrecht University. The project has two full-time post-doctoral researchers.
- Theme: Face:LIFT is investigating state of the art – especially adversarial – neural techniques to generate facial images from textual descriptions, in an interactive setting. An initial description leads to the generation of a facial image, which can be edited through additional linguistic input.

NLX4AI: Interactive Natural Language Technology for Explainable Artificial Intelligence 2020 to 2024

- Funder: Marie Skłodowska Curie (International Training Network)
- Collaborators: The project brings together 10 European universities and research centres, and an additional 10 industry partners.
- Roles: Beneficiary, principal supervisor (1 PhD project), co-supervisor (3 PhD projects), chair of the Training and Career Development Committee (TCDC).
- Theme: NL4XAI is training 11 Early-Stage Researchers to develop cutting-edge technology for generating natural language explanations of AI models. The research I conduct in NL4XAI focuses on model transparency for multimodal, Vision and Language neural models, and on explainable text generation models.

MUFINS: MUltilingual FInancial News Summarisation 2020 to 2022

- Funder: Malta Enterprise R&D Grant Scheme
- Collaborators: CityFalcon Trading (industrial partner).
- Roles: Principal Investigator (UM), leading a team of one post-doc and three senior researchers.
- Theme: MUFINS is an academic-industry collaboration involving UM and CityFalcon Trading Ltd, with the goal of providing bespoke summaries of multi-source financial news to corporate clients. The sub-project led by UM focused on multilingual neural models for learning language- and task-invariant representations.

Comparing and describing quantities 2022 to 2023

- Funder: Utrecht University Human-Centred AI Focus Area
- Collaborators: prof. dr. Stefan van der Stigchel, Psychology; prof. dr. Yoad Winter, Linguistics.
- Theme: This small project is an initial collaboration between computing, psychology and linguistics. Eye-tracking methods are used to understand how humans estimate quantities in visual scenes when prompted by linguistic stimuli. The findings will inform the development of benchmark tests to evaluate the sensitivity of image-to-text generation models to linguistic mass and count expressions.

MASRI: Maltese Automatic Speech Recognition   2019 to 2021

- Funder: University of Malta Research Excellence Awards
- Role: Principal Investigator, leading a team consisting of three senior researchers and a full-time post-doc.
- Theme: Creation of the first automatic speech recognition technology for Maltese, with particular focus on data curation and the use of transfer learning for ASR in low-resourced languages.

MAPA: Multilingual Anonymisation for Public Administrations   2019 to 2021

- Funder: European Commission under the Connecting Europe Facility (CEF)
- Collaborators: The project brings together 7 European industrial and academic partners.
- Role: Beneficiary, leader of the UM team consisting of two part-time research students and three senior researchers.
- Theme: Development of a multilingual Named Entity Recognition and anonymisation toolkit for all the languages of the European Union.

Assiduity: Intelligent Legal Search   2017 to 2019

- Funder: Take-Off Seed Fund Awards, University of Malta
- Role: Principal Investigator, leading a team consisting of an additional senior researcher and one part-time research student.
- Theme: Deep text analysis of legal text and development of mobile applications for intelligent search through legal documents. Assiduity developed the first mobile application for intelligent search for the legal profession in Malta.

Dizzjunarju Malti   2015 to 2019

- Funder: Malta Communications Authority and Vodafone Foundation
- Role: Project leader
- Theme: Creation of electronic lexical resources for Maltese, and dissemination for the general public.

MLRS: Maltese Language Resource Server   2007 to present

- Role: Principal Investigator
- Theme: Creation of computational resources, corpora and NLP tools for the languages of Malta, including large text corpora, part of speech taggers and tokenisers.

MAMCO: Maltese Multimodal Corpora   2011 to 2012, 2012 to 2013, 2015 to 2016

- Funder: University of Malta Research Fund
- Role: Co-investigator

- Theme: Collect of naturalistic dialogue data in Maltese, with semi-automatic annotation and speech and gestural level for machine learning of gestural mechanisms.

Morphological Productivity in Maltese: Corpus-based and experimental evidence 2010 to 2013

- Funder: University of Malta Research Fund
- Role: Principal Investigator
- Theme: Corpus-based, psycholinguistic/experimental and machine-learning methods for the assessment of productivity in Maltese word formation.

Bridging the Gap between Psycholinguistics and Computational Linguistics: The case of Referring Expressions 2010 to 2013

- Funder: NWO (Dutch National Science Foundation)
- Role: Postdoctoral Research Fellow (0.2 fte)
- Themes: Designing and implementing computational algorithms for referring expression generation, backed by psycholinguistic experiments on the production of such expressions by human speakers.

BabyTalk 2007 to 2010

- Funder: Engineering and Physical Sciences Research Council (EPSRC), UK
- Role: Postdoctoral Research Fellow (1.0 fte)
- Theme: BabyTalk was a highly influential project involving a collaboration between the Universities of Aberdeen and Edinburgh, and the Edinburgh Royal Infirmary (NHS, UK). It combined insights from Intelligent Signal Processing, Natural Language Generation, Pediatrics, Nursing and Cognitive Psychology to design systems to automatically summarise clinical neonatal intensive care data for decision support and information transfer for various user groups (nurses, doctors and family members).

TUNA: Towards a UNified Algorithm for the Generation of Referring Expressions 2004 to 2007

- Funder: Engineering and Physical Sciences Research Council (EPSRC), UK
- Role: PhD student (1.0 fte)
- Theme: Development of algorithms and datasets for referring expression generation. PhD thesis focused on content determination in the automatic generation of definite references to sets, combining experimental psycholinguistic techniques, corpus analysis and algorithmic modelling.

MaltiLEx: Maltese Lexical Resources 2001 to 2002

- Funder: University of Malta and Bank of Valletta
- Role: Research Assistant (1.0 fte)
- Theme: Construction of annotated language resources for Maltese, and the automatic extraction of lexical information from corpora.

Multi3Gen: Multi-task, Multilingual, Multimodal Language Generation
2019 to 2024

- Funder: COST Action CA18231
- Role: Management committee member; Working group leader (WG1: Grounded multimodal reasoning and generation) from 2019 to 2021.
- Aim: To foster an interdisciplinary network working on different aspects of language generation and multimodal natural language processing.

RefNet: Linking Research on Reference   2013 to 2016

- Funder: Engineering and Physical Sciences Research Council (EPSRC), UK
- Role: Participant, co-investigator
- Aim: Research network advancing collaboration between research communities working on reference

PARSEME: PARSing and Multiword Expressions   2013 to 2017

- Funder: COST Action IC1207
- Role: Management committee member
- Aim: COST action focusing on the role of lexical and phrasal information in enhancing parsing models.

Best paper award   2019
Best Practices for the Human Evaluation of Automatically Generated Text (with: C. van der Lee, E. van Miltenburg, S. Wubben, and E. Krahmer), *12th International Conference on Natural Language Generation (INLG'19)*, Tokyo, Japan.

Best reviewer award   2011
Twelfth European Workshop on Natural Language Generation. Association for Computational Linguistics. Athens, Greece.

Doctoral scholarship   2004 to 2007
Fully-funded PhD studentship by the Engineering and Physical Sciences Research Council (EPSRC), UK.

Conferences and workshops (selection)

- *Representation-Mediated Multimodality* 2022. Summer School organised as part of Multi3Generation COST Action, Schloss Etelsen, Germany.
- NLG in the Lowlands 2022: A workshop that brought together researchers in the Netherlands and neighbouring countries who work on all aspects of language generation.
- P-VLAM 2020 & 2022: Co-organiser, Workshop on People in Language, Vision and Mind, co-located with LREC 2020 and LREC 2022

- IntelLang 2020: Co-organiser, Workshop on Intelligent Information Processing and Natural Language Generation, co-located with ECAI 2020
- EMNLP 2019, 2020, 2021: Area chair, International Conference on Empirical Methods in Natural Language Processing
- INLG 2019, 2020, 2021, 2022: Area chair, International Conference on Natural Language Generation
- INLG 2018: Co-organiser and programme co-chair, 11th International Natural Language Generation Conference, Tilburg University, The Netherlands, November 2018.
- INLG 2017: Generation Challenges chair, 10th International Natural Language Generation Conference, Santiago de Compostela, Spain, 4-7 September, 2017.
- INLG 2016: Workshop chair, 9th International Natural Language Generation Conference, Edinburgh, UK, 5-8 September, 2016.
- ENLG 2015: Co-chair, Fifteenth European Workshop on Natural Language Generation, Brighton, UK, 10-11 September 2015.
- AMLaP 2015: Co-organiser, Conference on Architectures and Mechanisms for Language Processing, Valletta, Malta,3-5 September 2015.
- RefNet: Co-organiser, RefNet Summer school on Psychological and Computational Models of Language Production, School of Informatics, University of Edinburgh, 24-31 August 2014.
- COLING 2014: Programme Co-Chair (Generation and Summarisation track), 25th International Conference on Computational Linguistics, 23-30 August 2014, Dublin, Ireland.
- ENLG 2013: Co-chair, Fourteenth European Workshop on Natural Language Generation, Sofia, Bulgaria, in conjunction with the Annual Meeting of the Association for Computational Linguistics.
- Pre-CogSci 2009, 2011, 2013: Co-organiser, Workshops on Production of Referring Expressions: Bridging the Gap between Computational and Empirical Approaches, held in conjunction with the Annual Meetings of the Cognitive Science Society (Amsterdam 2009, Boston 2011, Berlin 2013)
- UCNLG+Eval 2011: Co-organiser, Workshop on Language Generation and Evaluation, held in conjunction with the 2011 Conference on Empirical Methods in Natural Language Processing (emnlp'11)

Shared tasks and evaluation initiatives

- WebNLG Shared Task 2022-23: This shared task, which is ongoing, is a collaboration between the University of Malta, CNRS/LORIA in Nancy, Dublin City University and Utrecht University. It has developed the first datasets for generation in under-resourced languages, namely Irish, Welsh, Breton and Maltese, and aims to evaluate the zero-shot multilingual transfer capacity of current neural encoder-decoder models. Results will be presented in 2023.
- Generation Challenges 2009-13: I co-founded the Generation Challenges in 2009 and was a member of the steering committee until 2013. This is an umbrella event bringing together a variety of Shared Task and Evaluation Challenges in Natural Language Generation and related areas (such as the development of dialogue systems and automatic summarisation), providing a platform for research development and comparable, replicable evaluation. The Challenges also serve as a launching pad for new proposals for Shared Tasks. The Challenges have been organised annually since 2009 and held in conjunction with the

European Workshop on Natural Language Generation (Athens, Greece, 2009; Nancy, France, 2011 and Sofia, Bulgaria, 2013) and the International Conference on Natural Language Generation (Salt Fork, Ohio, USA, 2010 and Starved Rock, Illinois, USA, 2012).

- TUNA-REG 2007, 2008 and 2009: Co-organiser, NLG Shared Task and Evaluation Challenges on Referring Expressions Generation. These were the first shared tasks organised in the NLG community.

Membership of professional societies

- Elected board member, SIGGEN: The ACL Special Interest Group in Natural Language Generation (2015 to 2019)
- Member, Association for Computational Linguistics
- Member, Cognitive Science Society (2009-2014)

Grant award committees

I have served as a member of evaluation committees for the following research funding agencies:
- ANR - French National Research Agency (2015, 2017)
- Research Foundation - Flanders: FWO Excellence Prize (2020)
- SNSF - Swiss National Sciecne Foundation (2022)

Candidate selection committees

I served as a member of the following committees to evaluate applicants for tenured or tenure-track positions:
- CiMEC - Centre for Mind/Brain Sciences, University of Trento, Italy. Member of the committee for job applications under the call *Bando RTD b) sc 09/H1 (Sistemi di elaborazione delle informazioni)* (March, 2022)
- Department of Information and Computing Sciences and Department of Pharmacology, Utrecht University. Member of the interview committee (teaching and education) for a joint cross-departmental position. (February 2022)

Journal reviewing

I have been invited to review for various top-tier journals in NLP, AI and psycholinguistics, including, among others: Transactions of the Association for Computational Linguistics (TACL); Computational Linguistics; Artificial Intelligence; Journal of Artificial Intelligence Research (JAIR); Computer Speech and Language; Natural Language Engineering; PLoS One; Journal of Experimental Psychology: Learning, Memory and Cognition; Discourse Processes; ACM Computing Surveys; Fuzzy Sets and Systems; IEEE Transactions on Computational Intelligence and AI in Games; Language Resources and Evaluation.

Conference program committees

I am annually invited to be a member of the program committee for international conferences in the fields of NLP, NLG and AI, including: ACL, EMNLP, NAACL, EACL, IJCAI, COLING, INLG and IJCNLP.

Conference/Workshop keynotes

- 7 May, 2023: *Describing scenes, actions and reasons: Some data, some models and quite a few explanations.* Keynote talk at the 9th International Symposium on Brain and Cognitive Science, Özyeğin University, Istanbul, Turkey.

- January, 2022: Third European Language Resource Coordination (ELRC) workshop, Malta

- April, 2022: NL4XAI Training Event on Natural Language Generation and Explanainability, LORIA and Universite de Lorraine, France

- November 2020 (postponed due to the pandemic): 2nd Annual EURNLP Summit, Paris, France

- September 2019: Artificial and Human Intelligence: Formal and Cognitive Foundations for Human-Centred Computing. Workshop at the European Conference on Artificial Intelligence (ECAI). Online

- September 2019: NLP4AI: 4th Workshop on Natural Language for Artificial Intelligence, Calabria, Italy

- August 2019: Python4NLP: Summer School in Natural Language Processing, Nancy, France

- October 2019: ICALP 2009: The 7th International Conference on Arabic Language Processing, LORIA and Universite de Lorraine, France

- August 2018: CNL 2018: The 6th International Workshop on Controlled Natural Languages, Co. Kildare, Ireland

- July 2014: *Symposium on Pragmatic Alternatives*, 36th Annual Meeting of the Cognitive Science Society, Quebec City, Canada

Invited seminars

- May 2022: Lecture for the Data Science and AI team at Ordina, The Netherlands

- October 2022: Computational Linguistics Seminar (CLS), Universiteit van Amsterdam, The Netherlands

- November 2021: NLP Group, University of Groningen, The Netherlands

- November 2021: ICS Colloquium, Dept. Information and Computing Sciences, Utrecht University, The Netherlands

- January 2018: Institut für Maschinelle Sprachverarbeitung (Institute for Natural Language Processing), Universität Stuttgart, Germany

- June 2017: Centro de Investigación en Tecnologías de la Información y de las Comunicaciones de la Universidad de Granada (CITIC-UGR), University of Granada, Spain

- December 2016: Centro Singular de Investigacion en Technoloxías de Informacion (CiTIUS), University of Santiago de Compostela, Spain

- February 2014: Copenhagen Linguistics Circle, University of Copenhagen, Denmark

- February 2014: Laboratoire d'Informatique de Grenoble, Grenoble University, France

- March 2011: Tilburg Center for Communication and Cognition, Tilburg University, The Netherlands

- June 2010: Computing Department, Open University, UK

- January 2010: LORIA, INRIA, Nancy, France.

- January 2010: Laboratoire d'Informatique de Grenoble, Grenoble University, France.
- December 2009: Malta Group of Professional Engineering Institutions, Malta.
- January 2009: Department of Computing Science, University of Aberdeen, UK.
- July 2008: School of Informatics, University of Edinburgh, UK.

SUPERVISION

PhD projects supervised to completion
- 2015-2019: Marc Tanti (principal supervisor). *Neural architectures for image captioning.* Institute of Linguistics and Language Technology, University of Malta.
- 2012-2016: Claudia Borg (principal supervisor). *Machine learning of morphology and unsupervised lexical clustering.* Institute of Linguistics and Language Technology, University of Malta.
- 2012-2016: Dr Luke Galea (co-supervisor). *Experimental account of syllabification and phonetic realisation.* Institut für Linguistik - Universität zu Köln, Germany.

PhD projects in progress
- 2022-present: Yingjin Song (principal supervisor). *Generating narrative text from image data.* Dept. of Information and Computing Sciences, Universiteit Utrecht.
- 2022-present: Eduardo Calò (co-supervisor). *Explaining logical formulas in natural language.* Dept. of Information and Computing Sciences, Universiteit Utrecht.
- 2022-present: Burak Kilic (promotor). *Classification and generation of legal text with large language models.* Dept. of Information and Computing Sciences, Universiteit Utrecht.
- 2020-present: Michele Cafagna (principal supervisor). *Multimodal semantic grounding and explainability in neural vision and language models.* Institute of Linguistics and Language Technology, University of Malta and Dept. of Information and Computing Sciences, Universiteit Utrecht.
- 2020-present: Ettore Mariotti (co-supervisor). *Explaining black-box models in terms of grey-box twin models.* CITIUS, University of Santiago de Compostela and University of Malta.
- 2020-present: Juliette Faille (co-supervisor). *Explainable models for text production.* CNRS, Université de Lorraine, France and University of Malta.
- 2019-present: Amanda Muscat (principal supervisor). *Stylistic features for trait and personality profiling in Maltese.* Institute of Linguistics and Language Technology, University of Malta.

Post-doctoral researchers
I hired and supervised the following post-doctoral researchers on past projects:
- 2020-2021: Marc Tanti *MUFINS: NLP-Driven FInancial News and content Search.*
- 2019-2020: Carlos Hernandez Mena. *MASRI: Maltese Automatic Speech Recognition.*

Masters theses
I regularly supervise and co-supervise MA and MSc theses in Linguistics, Natural Language Processing, Computer Science, AI and Psycholinguistics. At the University of Malta, I have supervised over 15 masters projects. At Utrecht University, I am daily supervisor and first examiner of three projects in the MSc Artificial Intelligence program.

PhD examiner roles (selection)

- 2023: External committee member ('pre-examiner'), Leo Leppänen, *Methods for Automated Generation of Natural Language Reports*, University of Helsinki, Finland.

- 2023: Examiner, Sonya Nikiforova, *Image Captioning with External Knowledge*, Institute for Language Sciences, Utrecht University, The Netherlands.

- 2023: External examiner, Chris Emmery, *User-Centered Security in Natural Language*, Tilburg University, The Netherlands.

- 2020: Internal examiner, Marthese Borg, *The Curious Case of Lexical Ambiguity: Using Talker-Specific Characteristics and Facial Expressions to Decode the Puzzle*. Department of Cognitive Science, University of Malta.

- 2019: External examiner, Ravi Shekhar, *Learning to merge language and vision: A deep evaluation of the encoder, the role of the two modalities, the role of the training task*. Department of Information Engineering and Computer Science, Universitá degli Studi di Trento, Italy.

- 2019: External examiner, Lindley Winchester, *The Morphosyntax of the Maltese DP*, Department of Linguistics, Georgetown University, Washington DC, USA.

- 2018: External examiner, Thiago Castro-Fereira, *Advances in Natural Language Generation*, Tilburg University, The Netherlands.

- 2017: External examiner, Hadi Banaee, *From Numerical Sensor Data to Semantic Representations: A Data-Driven Approach for Generating Linguistic Descriptions*, Department of Computer Science, Örebro University, Sweden.

- 2016: External examiner, Alejandro Ramos-Soto, *Application of Fuzzy Sets in Data-to-Text Systems*, Centro Singular de Investigación en Tecnoloxías da Información, University of Santiago de Compostela, Spain.

- 2016: External examiner, Bikash Gyawali, *Surface Realisation from Knowledge Bases*, LORIA, Université de Lorraine, Nancy, France.

- 2015: External examiner, Marieke Hoetjes, *Talking Hands: Reference in Speech, Gesture and Sign*, Tilburg University, The Netherlands.

- 2015: External examiner, Bikash Gyawali, *Surface realisation from knowledge bases*, Université de Lorraine, France.

- 2014: External examiner, Jorrig Vogels, *Referential Choices in Language Production*, Tilburg University, The Netherlands.

- 2009: External examiner, Carlos Leon Aznar, *A computational model for automated extraction of structural schemas from simple narrative plots*. Department of Computer Science, Universidad Complutense de Madrid, Spain.

MUFINS: Consultancy and collaboration 2020

- Partners: CityFalcon Trading Ltd and Malta Enterprise
- Role: As part of the MUFINS project, in addition to cutting-edge research on neural solutions to domain and language-invariance, we offer consultancy to our industrial partners to help them extend their software services to a broad range of international clients, in multiple languages.

Software consultant 2012

- Company: Brighthub (USA)
- Role: Consultancy on design and implementation of software to generate academic reports.

Software consultant 2011

- Company: Agfa Healthcare (USA)
- Role: Design and implementation of a software client to generate medical reports from patient data.

Manuscript consultant 2008

- Company: Springer Publishers Ltd (Heidelberg)
- Role: Manuscript reviewer and professional consultant

Lexicographer 2007

- Company: Merlin Publushers Ltd (Malta)
- Role: Consultant on use of computational lexicographic techniques for the compilation of a children's dictionary.

ICT Committee, National Council for the Maltese Language 2014 to 2019
The National Council for the Maltese Language nominated me to chair the Committee on Human Language Technology. As part of my work on this committee, I launched a policy proposal on NLP and the curation of digital language resources for the languages of Malta.

European Researchers' Night 2015 to 2021
As Director of the Institute of Linguistics and Language Technology, I ensured that the Institute contributed annually to the events organised as part of the annual Science in the City in Valletta. Our contribution involves interactive activities showcasing research in NLP.

Café Scientifique 2014
Tales from your Computer: Invited public lecture at the Malta Café Scientifique on Artificial Intelligence for automatic text generation.

Popular science writings and media

I regularly contribute to publications intended for a broad audience, with a view to sensitising the public about scientific research and results. The following is a selection of such publications and other media contributions.

Gatt, A. (2020). *On the generation of language*. In: *Newspoint*.

Camilleri, C., A. Gatt, and G. Pace (2018). *Smart search for Maltese legal professionals*. In: *Think Magazine*.

Gatt, A. (2018b). *Drittijiet diġitali u ugwaljanza lingwistika [Digital rights and language equality]*. In: *Encore - Arts and Culture Magazine* 13.

Gatt, A. and G. Pace (2018). *Contracts at the crossroads*. In: *Sunday Times of Malta*.

Gatt, A. (2017a). *Is a face worth a thousand words?* In: *Sunday Times of Malta*.

– (2017b). *Laboratorju tal-Lingwi: Il-Malti, it-teknoloġija u l-mezzi soċjali [A language lab: Maltese, technology and social media]*. In: *Lehen il-Malti* 36.

Jien min Jien? Il-Lingwa. [Who am I? Language and the bigger picture] (2015). Documentary written and produced by the Science in the City / European Researchers Night team led by Albert Gatt. Directed by Manuel Ciantar. Malta: Studio Seven and Public Broadcasting Services.

Gatt, A., G. Pace, and M. Rosner (2014). *Decoding language*. In: *Think Magazine* 8.

Gatt, A. (2013). *Il-Malti, ir-riżorsi diġitali, u t-traduzzjoni [Maltese, digital resources and translation]*. In: *L-Aċċent: Journal of the Maltese Language Department of the European Commission* 10.

IMPACT: SOFTWARE AND RESOURCES

BERTu Model and Data
Pretrained and finetuned BERT models and 500m word text corpus. This is the first large-scale initiative to train deep transformer networks for NLP in Maltese, an under-resourced language.

SimpleNLG
Java and Python API for Natural Language Generation. Used by a large number of research and commercial teams and now ported to several other languages, including Portugese, German, and Galician.

MLRS
Maltese Language Resource Server: An online collection of electronic language resources, including the *Korpus Malti*, a 250+ million token corpus of pos-tagged and lemmatised Maltese text. Other resources include online lexica. Since its launch, the corpus is being used by over 1000 users across four continents.

GenChal Repository
A repository of legacy evaluation data from several editions of the Generation Challenges related to different sub-tasks of Natural Language Generation, together with software to compute evaluation metrics. The data is coupled with evaluation results from various shared tasks.

TUNA Corpus of referring expressions
Semantically annotated corpus of references to objects in English. This corpus formed the basis of a number of international NLG shared tasks, and is still widely in use for

research purposes.

Maltese Dictionary
Online full-form Maltese-English dictionary constructed semi-automatically, with web service API for integration with NLP applications, as well as public use. This resource is also widely used by members of the public, including educational institutions.

Maltese Sign Language
Online multimodal lexicon on Maltese Sign Language (LSM), constructed in collaboration with the LSM research group at the University of Malta and the Maltese Deaf Community.

PUBLICATIONS

Cited over > 4900 times. Google $h-$index: 32; $i10-$index: 64.

A full publication list with downloadable papers is available on my homepage, at https://albertgatt.github.io.